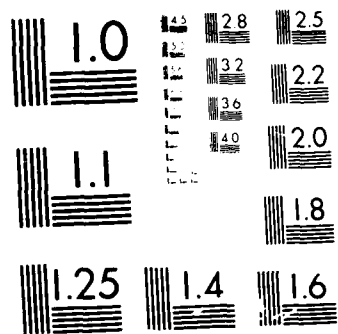


AD-A194 716 DYNAMICALLY RECONFIGURABLE DISTRIBUTED DATABASE SYSTEMS 1/1  
(U) CONNECTICUT UNIV STORRS F HARYANSKI MAR 88  
ARO-21826.15-EL DAAG29-85-K-0044

UNCLASSIFIED

F/G 12/7 NL





## REPORT DOCUMENTATION PAGE

AD-A194 716

SELECTED

MAY 17 1988

1b. RESTRICTIVE MARKINGS

DTIC FILE COPY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

3. DISTRIBUTION/AVAILABILITY OF REPORT

Approved for public release;  
distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

ARO 21826.15-EL

6a. NAME OF PERFORMING ORGANIZATION

Univ. of Connecticut

6b. OFFICE SYMBOL

(If applicable)

7a. NAME OF MONITORING ORGANIZATION

U. S. Army Research Office

6c. ADDRESS (City, State, and ZIP Code)

Storrs, CT 06268

7b. ADDRESS (City, State, and ZIP Code)

P. O. Box 12211

Research Triangle Park, NC 27709-2211

8a. NAME OF FUNDING/SPONSORING ORGANIZATION

U. S. Army Research Office

8b. OFFICE SYMBOL

(If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

DAAG29-85-K-0044

8c. ADDRESS (City, State, and ZIP Code)

P. O. Box 12211

Research Triangle Park, NC 27709-2211

10. SOURCE OF FUNDING NUMBERS

PROGRAM  
ELEMENT NO.PROJECT  
NO.TASK  
NO.WORK UNIT  
ACCESSION NO.

11. TITLE (Include Security Classification)

Dynamically Reconfigurable Distributed Database Systems

12. PERSONAL AUTHOR(S)

Fred Maryanski

13a. TYPE OF REPORT

Final

13b. TIME COVERED

FROM 6/28/85 TO 1/27/88

14. DATE OF REPORT (Year, Month, Day)

March 1988

15. PAGE COUNT

8

16. SUPPLEMENTARY NOTATION

The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

>Database Systems; Network Topology; Topology,  
Network Nodes; System Design Strategy.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

One of the few common threads throughout most distributed database research projects has been the assumption that the network topology is stable. However, mobility and portability are often essential features of network nodes, especially (but not exclusively) in military applications. This project investigated both basic and applied research problems which arise when the presupposition of fixed topology is removed. The goal of this effort was to explore the essential issues of dynamically reconfigurable database systems and to develop a design strategy for such systems. This report contains the results of this investigation.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT

☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION

Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL

22b. TELEPHONE (Include Area Code)

22c. OFFICE SYMBOL

DYNAMICALLY RECONFIGURABLE DISTRIBUTED DATABASE SYSTEMS

Fred Maryanski

March 1988

U.S ARMY RESEARCH OFFICE

GRANT NUMBER DAAG 29-85-K-0044

FINAL REPORT

University of Connecticut

APPROVED FOR PUBLIC RELEASE;

DISTRIBUTION UNLIMITED

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

## 1. PROBLEM STATEMENT

### 1.1. Overview

One of the few common threads throughout most distributed database research projects has been the assumption that the network topology is stable. However, mobility and portability are often essential features of network nodes, especially (but not exclusively) in military applications. This project investigated both basic and applied research problems which arise when the presupposition of fixed topology is removed. The goal of this effort was to explore the essential issues of dynamically reconfigurable database systems and to develop a design strategy for such systems.

### 1.2. General System Characteristics

This study was targeted toward networks with the following characteristics:

- (1) The network consists of two types of nodes - servers and clients. The servers offer specialized functionality which may be accessed by any authorized node.
- (2) Access to a server may be initiated by any node (client or another server). Access to a client may, in general, be initiated only by a server.
- (3) The database management facilities will be provided by a database server. Other servers will provide functions such as naming, authentication, and network monitoring.
- (4) Any node may disappear from the network (become unreachable) at any time with the expectation that it will reappear after an unspecified period of time, possibly at a new physical location.
- (5) A link between any two nodes is transient even while both nodes are connected to the network and lasts only for the duration of a session (initiated by one of the nodes).
- (6) The two nodes engaged in a session may or may not be reachable by other nodes during the session, depending on their level of sophistication (e.g. number of dial in ports).
- (7) The majority of network entities are highly mobile client nodes which occasionally appear, interact with the network, and then depart.

Dist	Avail and/or Special
A-1	

- (8) Server nodes are also mobile; however they will tend to be more stable than the client nodes.
- (9) Though each server is a separate and unique logical entity, each may be implemented on one or more physical nodes, and a single node may offer multiple services.

At the highest level, the network consists of a collection of mobile nodes which communicate on dynamically established links. The database and other major network facilities are contained on a group of server nodes which, although mobile, have a high degree of availability. As a consequence, server functionality will be replicated. Client nodes appear at different remote locations, attach to a server for the purposes of either adding information to the database or querying its contents, and then depart. In some situations, clients may pose a query, depart, and then return at another location to request the result. The remote client nodes may exchange messages through a post office, but there is no direct client communication during a session.

### **1.3. Objectives - Scope and Issues**

The dynamic nature of a system as outlined above imposes special constraints on the designer. Solutions and methods that apply for distributed database systems that are built on top of static or quasi-static communication networks do not, in general, apply for a system built on top of a highly dynamic network. It was therefore necessary to investigate the set of research issues listed below to determine the impact of a dynamic topology:

- data allocation,
- query optimization,
- concurrency control,
- survivability,
- network management,
- communication protocols,
- routing,
- simulation methodology.

## **2. RESULTS**

### **2.1. Data allocation**

A number of dynamic data allocation strategies for dynamically reconfigurable distributed databases have been developed. The primary tenet underlying the data allocation strategy is that reallocation should take place only when overall system performance would be improved or if the number of copies of a file falls below a pre-specified minimum. The allocation algorithms choose the best possible assignment using heuristic benefit functions and greedy search strategies. The goal of the technique is to select an assignment as close as possible to optimal while minimizing search time. Performance experiments indicated that the allocations chosen by these algorithms were on the average within 2% of the optimal assignment while using 3 orders of magnitude less CPU time. These results are reported in detail in [CP-4]. The overall performance of the algorithms was then enhanced by the development of techniques for the parallel reorganization of the database when reallocation is necessary, see [CP-2].

### **2.2. Query optimization**

Two query optimization techniques developed for static distributed database systems were modified to consider the problems of a dynamic environment. The strategies were the dynamic execution semijoin approach of Yu and Chang and the bucket semijoin technique of Krishnamurthy and Morgan. Experimental results indicated that while the dynamic nature of the network impacted the performance of both strategies, the dynamic execution approach had a smaller increase in running time. Details of this study appear in [TR 6,T-3].

### **2.3. Concurrency control**

The methodology developed for concurrency control in a dynamically reconfigurable environment is the Enhanced Multiversion Timestamp ( EMT ) approach [T-5] which is an extension of Reed's timestamping algorithm. EMT contains a special VIEW operation which permits reading of data without regard for locks, handles read only transactions specially, and attempts to avoid redoing write operations whenever possible. Performance experiments indicate that the above enhancement improve both the average and worst case

behavior of transactions operating under this currency control approach as compared with basic timestamping.

#### **2.4. Survivability**

The database management system must be able to withstand the departure of any number of client nodes and any single server node. Initial work in this area was predicated on the assumption that the underlying communication media would preclude network partitioning. This assumption is valid in many networks such as common-carrier phone links. In this context, algorithms for reliable transaction management and commit were developed to handle both planned and unexpected departures of server nodes [CP-7]. The transaction management algorithm can tolerate up to  $k$  simultaneous failures, providing that  $k$  is less than the total number of database server nodes. The commit protocol requires only the existence of one server for each data object in the read or write set at the appropriate time in the protocol.

The problem of partitioning was treated by a different set of algorithms [CP-1,CP-3] which made an initial assumption of a hierarchical topology. The algorithms provide take an optimistic approach to the partitioning problem, allowing work in individual partitions to proceed as far as possible and then reconciling differences at reconnection. The partitioning algorithms are unique in that they permit nodes to reattach at points other than their point of departure, thus supporting complete reconfiguration. The placement of network monitors is a key element of the partition handling strategy. An algorithm has been developed for the optimal placement of network monitors in a hierarchical network.

#### **2.5. Network Management**

The responsibility for the management of the network resides with the Network Monitor and Name Servers. The design and implementation of the network management facility is report in [TR-5]. These servers provide for both planned and unexpected departures from and returns to the network.

#### **2.6. Communication Protocols**

A network communication facility was implemented on top of the Unix interprocess communication subsystem. Protecs for establishing and terminating links and ensuring reliable, error-free, efficient



communication. Details of these link level protocols are provided in [T-4].

In the communication model followed in this project, a session layer mediates between the low level communication links and the high-level applications. The protocols developed [T-1] enable processes to return to and depart from the network, to establish sessions between any two processes, and to send and receive messages.

## **2.7. Routing**

The highly dynamic nature of the network considered necessitated the development of a new approach to routing. The strategy followed involved the enhancement of Floyd's basic path algorithm with alternate routing procedures [T-2] which seeks new subpaths when a topological change effects a given route. Simulations indicate that this approach produces substantial performance benefits in a network with a dynamic topology.

## **2.8. Simulation Methodology**

Simulation has proven to be a valuable tool in the projection of the performance of many of the new algorithms developed during this project [CP-6]. During the course of these simulation experiments, a number of shortcomings of existing simulation techniques were noted. Consequently, a simulation package oriented toward the simulation of distributed systems has been developed [T-3-1]. Among the special features of this package are portability, interfaces special purpose simulation language with C, direct interface to operating system, extensibility, special constructs to support simulation of distributed systems, generation of multiple plots, interactive debugger, control of statistics produced.

# **3. PUBLICATIONS AND TECHNICAL REPORTS**

## **3.1. Conference Publications**

CP-1. Wong, D. and Maryanski, F., "Simulation of Partition Processing Strategies for Driftwood", *Summer Simulation Conference*, July 1988.

- CP-2. Du, X. and Maryanski, F., "Data Reorganization in a Dynamically Reconfigurable Environment", *International Conference on Distributed Computing Systems*, June 1988.
- CP-3. Wong, D., "Transaction Processing in Partitioned Dynamically Reconfigurable Distributed Database Systems", *ACM SIGSMALL/PC Conference*, May 1988.
- CP-4. Du, X. and Maryanski, F., "Data Allocation in a Dynamically Reconfigurable Environment", *International Conference on Data Engineering*, Feb. 1988, pp. 74-81.
- CP-5. Kemeny, J., Lambert, D., and Maryanski, F., "Performance Projections for Relational Query Processors", *International Workshop on Database Machines*, Oct. 1987, pp. 513-531.
- CP-6. Rubinovitz, H. and Maryanski, F., "Performance Modeling in Dynamically Reconfigurable Distributed Database Systems", *Pittsburgh Conference on Modeling and Simulation*, Apr. 1987.
- CP-7. Zajac, S. and Maryanski, F., "Optimistic Concurrency Control and Reliable Commit in Dynamically Reconfigurable Distributed Databases", *ACM SIGSMALL/PC Conference*, Dec. 1986, pp. 137-145.
- CP-8. Kantorowitz, E., Maryanski, F., and Shasha, D., "Distributed Office-By-Example ( DOBE )", *International Conf. on Data Engineering*, Feb. 1986, pp. 166-174.

### 3.2. TECHNICAL REPORTS

All of these reports have been submitted for publication.

- TR-1. Rubinovitz, H. and Maryanski, F., "A Software Tool for Distributed Database Simulation", Feb. 1988.
- TR-2. Maryanski, F. and Kemeny, J., "Performance of Relational Query Processing Systems", Dec. 1987.
- TR-3. Maryanski, F., et al., "DRIFTWOOD: A Dynamically Reconfigurable Distributed Database System", Nov. 1987.

TR-4. Wong, D. and Maryanski, F., "Reliable Transaction Processing in a Partitioned Dynamically Reconfigurable Distributed Database", June 1987.

TR-5. Rubinovitz, H., Maryanski, F., and Gold, Y., "Network Services for Dynamically Reconfigurable Databases", Mar. 1986.

TR-6. Iyengar, S.V. and Maryanski, F., "Query Optimization in Dynamically Reconfigurable Distributed Databases".

TR-7. Gold, Y. and Moran, S., "A Correction Algorithm for Token-Passing Sequence in Mobile Communication Networks".

TR-8. Gold, Y. and Chen, M.Y., "Burst Noise and Bit-Error Rate in Direct-Sequence Spread-Spectrum Communication Systems".

TR-9. Gold, Y. and Kumritz, S., "Analysis of Busy-Tone and Busy-Message Access Protocols for Mobile, Multihop Networks with Channel Reuse".

TR-10. Chlamtac, I., Elazar, A., and Gold, Y., "Distributed Algorithm for Shortest Path Routing in Dynamic Networks".

### 3.3. THESES

The following theses were completed by students employed on the project.

T-1. Tran, T.T., "Design and Implementation of Session Establishment in Mobile Network", M.S. Thesis, Mar. 1988.

T-2. Nagy, S.C., "Routing in a Dynamically Reconfigurable Distributed Database System", M.S. Thesis, June 1987.

T-3. Iyengar, S.V., "Query Optimization in Dynamically Reconfigurable Distributed Database Systems", M.S. Thesis, Dec. 1986.

T 4. Lin, S.L.G., "Data Communication in Dynamically Reconfigurable Distributed Database Systems", M.S. Thesis, Nov. 1986.

T 5. Bittner, G.W., "The EMT Concurrency Control System", M.S. Thesis, Sept. 1986.

#### **4. PARTICIPATING SCIENTIFIC PERSONNEL**

The following individuals received partial research support from this project.

##### **4.1. Faculty**

- (1) Fred Maryanski, Principal Investigator
- (2) Yaron Gold, Investigator
- (3) Lester Lipsky, Investigator

##### **4.2. Graduate Research Assistants**

- (1) George Bittner, Completed M.S., Sept. 1986.
- (2) Xiaolin Du
- (3) Sridhar Iyengar, Completed M.S., Dec. 1986.
- (4) Gail Lin, Completed M.S., Nov. 1986.
- (5) Susan Nagy, Completed M.S., June 1987.
- (6) Harvey Rubinovitz
- (7) Tho Tran, Completed M.S., Mar. 1988.
- (8) David Wong
- (9) Sue Zajac

END  
DATE  
FILMED  
8-88  
DTIC